

What is claimed is:

- 1 1. A method of guiding a user in iteratively deriving object models from documents such
 2 as requirements documents and validating such object models against documents,
 3 comprising the following steps, which may be applied iteratively and interleaved in
 4 any order:
 - 5 a) identifying model elements using parts of speech and frequencies of
 6 word base forms and noun phrases in a document;
 - 7 b) establishing associations between the model elements using collocations
 8 and textual contexts of the word base forms and noun phrases
 9 corresponding to model elements in the document;
 - 10 c) validating object models using collocations and frequencies of word
 11 baseforms and noun phrases in the document, as well as natural language
 12 paraphrases of the models.
- 1 2. The method of claim 1, in which step (a) comprises the steps of:
 - 2 a) identifying classes using noun base forms and noun phrases frequently
 3 occurring in the document;
 - 4 b) identifying attributes using adjective base forms frequently occurring in
 5 the document;
 - 6 c) identifying associations between classes using verb base forms
 7 frequently occurring in the document.
- 1 3. The method of claim 1, in which the identification in step (a) is established by
 2 automatic linguistic processing of the document.
- 1 4. The method of claim 1, in which the model elements of step (a) are based on the
 2 concepts and notation of the Unified Modeling Language for representing object
 3 models.

- 4 5. The method of claim 1, in which the model elements of step (a) are based on the
5 concepts and notation of Entity-Relationship models.
- 6 6. The method of claim 1, in which step (b) comprises the steps of:
- 7 a) declaring associations between classes using collocations and textual
8 contexts of word base forms corresponding to the model elements in the
9 document;
- 10 b) associating attributes with classes using collocations and textual contexts
11 of the word base forms corresponding to the model elements in the
12 document;
- 1 7. The method of claim 1, in which the collocations and textual contexts are established by
2 automatic linguistic processing.
- 1 8. The method of claim 1, in which associations between the model elements of step (b)
2 are based on the concepts and notation of the Unified Modeling Language for
3 representing object models.
- 4 9. The method of claim 1, in which the model elements of step (b) and associations
5 between the elements are based on the concepts and notation of Entity-
6 Relationship models.
- 7 10. The method of claim 1, in which step (c) comprises the steps of:
- 8 a) detecting any missing model elements having corresponding word base
9 forms and noun phrases that occur with high frequency in the document;
- 10 b) detecting any model elements with corresponding word base forms and
11 noun phrases that occur with low or zero frequency in the document;
- 12 c) detecting any missing associations between classes or between classes
13 and their attributes corresponding to word base forms or noun phrase forms
14 that collocate in the document;

- 15 d) verifying the semantics of the model using descriptive paraphrases in
16 natural language.
- 1 11. The method of claim 1, in which the natural language paraphrases in step (c) are
2 automatically produced.